DOCUMENT RESUME

ED 397 393 CS 012 533

AUTHOR Macdonald-Ross, Michael; Scott, Bernard
TITLE The Readability of OU Foundation Courses. Text & Readers Programme Technical Report No. 5.
INSTITUTION Open Univ., Walton, Bletchley, Bucks (England). Inst. of Educational Technology.
PUB DATE 96
NOTE 17p.; A separate two-page summary has been appended.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Content Analysis; Higher Education; *Introductory Courses; *Readability; Reading Research; Textbook Content
IDENTIFIERS Flesch Reading Ease Score; *Open University (Great Britain)

ABSTRACT

The course units which form the core of the four foundation courses at the Open University were sampled for readability, using an electronic version of R. Flesch's (1951) Reading Ease Score. The textload (reading workload) of the courses was predicted by estimating the total length in words of all text components of these courses. Large variations in readability occurred between authors, topics, and courses. Large variations also occurred in the overall textload on these courses. The single best predictor of the reading difficulty of a course unit is the identity of the author, which suggests that one route to better readability would lie in the improved management of human resources in a course team. A substantial amount of the prose was more difficult than passage 3 on the original cloze reading test: prose difficulty is therefore likely to be a major factor in hindering course completion for students with weaker reading skills. (Contains 16 references, and 7 charts and figures of data. Four appendixes present data for each of the 4 foundation courses.) (Author/RS)

Text & Readers Programme

Technical Report #5

The readability of OU foundation courses

Michael Macdonald-Ross
Bernard Scott

Institute of Educational Technology
The Open University
Milton Keynes MK7 6AA

The readability of OU foundation courses

Abstract

The course units which form the core of four foundation courses at the Open University were sampled for readability, using an electronic version of Flesch's Reading Ease Score. The textload (reading workload) of the courses was predicted by estimating the total length in words of all the text components of these courses. Large variations in readability occur between authors, topics, and courses. Large variations also occur in the overall textload on these courses. The single best predictor of the reading difficulty of a course unit is the identity of the author, which suggests that one route to better readability would lie in the improved management of human resources in a course team.

A substantial amount of the prose is more difficult than passage 3 on our original cloze reading test: prose difficulty is therefore likely to be a major factor in hindering course completion for students with weaker reading skills.

Introduction

The work reported here is part of an investigation into the reading difficulties faced by students entering the OU at foundation course level. This project has two main parts. The first part, an investigation of OU students' reading skills, has been the subject of three previous reports (Macdonald-Ross & Scott 1995a, b; Scott & Macdonald-Ross 1995). This report deals with the readability of OU foundation course texts.

The total text length and readability of four foundation courses was estimated. It is now clear that much foundation course material is linguistically difficult for many of the students. An analysis of the third level history course A317 revealed a similar pattern.

The work involves certain novel features: it is one of the largest surveys of British educational learning material, and probably the largest survey of the reading difficulty of higher educational material in any country. The existence of printed teaching material, and particularly text in electronic form, made possible a much more extensive sampling than would have been possible by hand. Lastly, the data can be related to the known reading capacity of the audience, as measured earlier by our reading tests. This has rarely (if ever) been achieved for such a large group of learners in higher education.

Michael Macdonald-Ross
Bernard Scott

Acknowledgements

Clive Lawless for data on A317. **John Close** for three months work sampling readability. John was supported by a Quality Assurance grant, for which we are most grateful.

The rationale

One thing that distinguishes this research from most other research projects on our students is its concentration on psychological aspects of student learning. We are attempting to test a model of the effect of reading mismatch on student learning and progress.

In our model (figure 1) we suppose that mismatch between the reading skills of the students and the readability of the course materials is an important cause of student drop-out and failure. We compare our data on student reading skills and text readability levels to data on the progress of our cohort throughout the year. So far the model has worked well.

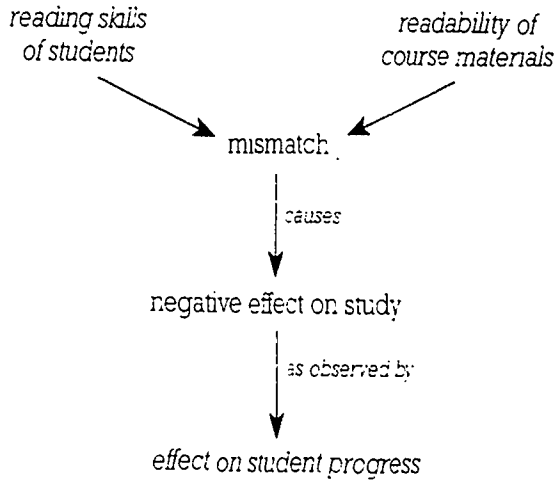


Figure 1. A model of the central hypothesis
A facet not captured by this graphic is the variety of other factors which also impact on student progress. The standing of reading mismatch relative to the other known factors in student progress is being investigated

Readability measures

Readability measures of the type we used are indices that predict: they do not explain how students learn, or teach authors how to write. They do not measure the conceptual difficulty of the subject-matter, unless indirectly this shows itself in more difficult prose.

What they do – and do well – is to predict the difficulty of prose, especially if there is some information about the reading skills of the readers.

In this report, then, the readability scores quoted are measures of syntactic and semantic complexity that predict the difficulty of text comprehension. This operational definition does not deny that many other factors make up a text, nor that investigation into learnability, conceptual comprehension, prose style and the effect of broader study skills may be relevant to a full appreciation of reader-text interaction.

Readability measures have been criticised, usually by authors who fail to discriminate between an index and an explanation, or between the production of prose and its predicted effect on the reader.

Our view on the matter is that, providing their limitations are well understood, there are few metrics in the whole of the social sciences which have been as extensively validated as the readability measures we use: their effectiveness at prediction has been well demonstrated.

Most formulæ enshrine a balance between ease of application and coverage of linguistic variables. Thus, good prediction may be got by two-variable formulæ where one variable is an index of semantic complexity (usually a direct or indirect measure of the vocabulary demands of the text), and the other variable is an index of syntactic complexity (usually sentence length).

These two indices combined are highly predictive of prose difficulty as experienced by readers. "The good news is that two language variables, and simple ones at that, account for so much of the variance in the prediction of readable writing" (Klare 1984). More complex formulæ developed for special purposes (such as linguistic research) add little extra to the accuracy of prediction.

Because of the extent of the literature on readability, it is not a simple matter to give an overview of the field. There are hundreds of studies on the effect of readability on various age groups, and of its effect under various conditions of motivation, prior knowledge of subject-matter, and its relation to the reading skills of subjects. The main resources to gain access to this literature are the reviews by George Klare (1963, 1974 and 1984). The topic was also addressed by Macdonald-Ross (1979), and by Hirsch in his *Philosophy of Composition* (1977).

A few examples: Klare and Smart (1973) found a close relationship between readability level of correspondence course material and the probability that correspondence students would persevere and complete their courses. Murphy (1947a,b) and Swanson (1948) found that readership of newspapers goes up with improved readability in split-run studies of newspapers. Sticht and colleagues showed that readability had a significant effect on the use of military manuals (Kern, Sticht & Fox 1970), and Johnson et al (1972) showed the effect of readability on discrepancies in following military procedures.

These results suggest that readability measures are as applicable to adults as to younger age groups. Indeed, the two best-known formulæ, the Flesch Reading Ease Score and the Dale-Chall formula, were actually developed for adult materials and validated against standard criteria for adult readers.

Textload

We use the term textload for that part of the coursework which is textual. The readability of texts has an effect on study which varies according to the reading skills of the students. The magnitude of this effect is influenced by the quantity of text and whether there is time in the schedule for re-reading and decoding meaning. It has been shown by Klare in a series of studies that the effect of a mismatch between reading skills and text readability is more pronounced if subjects are put under time pressure; learner motivation and familiarity and interest in the specific subject-matter are other relevant factors.

Assuming learner motivation is fairly high throughout our foundation courses (there is much informal evidence to suppose so), variation in course length remains as a significant factor. The reason for this is obvious: the heavier the workload, the more the student is under pressure to keep up with the work. The most significant element in course workload is the textual component. Our students are adults, often with jobs and families, which means that the pressure of workload is real, and may become critical.

Our system of continuous assessment means that students are driven by a schedule which cannot be easily evaded if they wish to succeed on the course. Tactics such as selective reading are no doubt often used (here we suffer from a

lack of good diary or observational data) but again the nature of our courses means that there is not so much incidental material as there might be in a traditional university course.

We would expect to see a greater effect of readability on overloaded courses and a lesser effect on underloaded courses (if any of the latter exist!). After its first few years the University has made great efforts to keep the students' workload within bounds. Courses are expected to contain about twelve hours work per week, and survey feedback is used to recover rough data on the time taken by students for study. There is little data on the total quantity of text in words, partly because we have only recently held the course text electronically, and partly because each course has a large number of different text elements.

We can get easily the total word count for the course unit material for any course which is held electronically, but all courses have a variety of booklets, set books, readers and guides, which raise the total word count. To give some idea, here are the overall figures for two foundation courses [estimated by sampling where electronic text was not available]:

Social Science D103:	course units	574,000
	other	323,550
	total:	897,550
Arts A102:	course units	512,200
	other	663,100
	total:	1,175,300

A clear pattern emerges here: A102 is slightly shorter than D103 if only the course units are considered, but the other course text in A102 is over twice as long as the other course text in D103. Such a difference is probably significant in its impact on student workload.

According to Ronald Carver's important and comprehensive review of the speed of reading (1990) "as long as college students and adults engage their rauding process they tend to read at a relatively constant rate around 300 words per minute even when the materials vary drastically in terms of difficulty level". Carver's rauding [reading for meaning] is his third gear. Lower gears are for learning and are correspondingly slower. Note that most units contain Activities and SAQ's [self-assessment questions] Considering all this, a mean reading speed for comprehension can hardly be more than 150 wpm, and might well be less.

Let us do the calculations for D103.

32 weeks x 12 hours = 384 hours study time

897550/150 = 5983.3 minutes or about 100 hours [99.72] to read text materials once.

Allowing 2 hours per tv programme, 30 minutes per radio programme and 2 hours for each 1 hour tape gives 54 hours for other media. 384–54 = 330 hours.

The Course Guide says about half of study time is for Course Units (and set book and reader?). This is about 190 hours. This leaves 140 hours for writing TMAs [tutor-marked assignments], though it is difficult to separate reading and writing aspects of study. For 8 TMA's, this is 17.5 hours per TMA [perhaps an underestimate].

It looks as if a student who studied 12 hours per week, of which an average of 6 hours was spent reading printed course materials, could hope to read everything in D103 once, and most of it twice. That is probably about right, so we conclude that D103 is not overloaded (in that respect, a triumph of course management).

By comparison, A102 contains significantly more text than D103, due entirely to the non-course unit material. If D103 is about right, then A102 is almost certainly overloaded, a factor which will lead to selective reading by some students and expenditure of extra time by others. One can be fairly certain that students with borderline reading skill will find it a difficult course: the pressure on their time will make the text readability a critical factor.

We clearly need empirical data from observational and diary studies to know how students cope with these difficulties. Such data is not available at present, and is not by any means easy to obtain in a distance learning system.

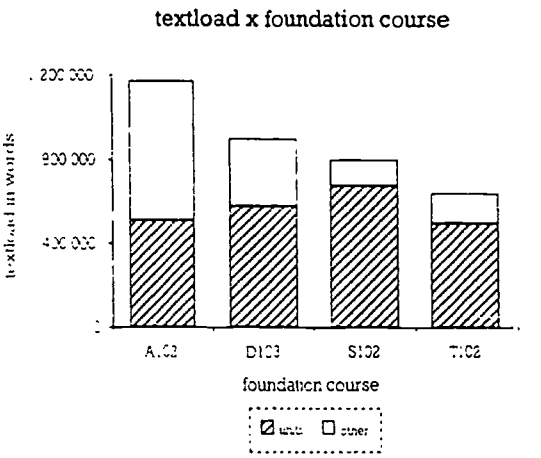


Figure 2 The textload of four foundation courses. A102 (Arts) has the greatest textload, and T102 (Technology) the least. Since all courses have various non-text activities, one cannot reach conclusions about overall course workload from this data alone

Note that S102 has the most text in the form of course units and little other text. Neither S102 nor T102 use set books or readers on these courses.

Sampling

Before the sampling a certain amount of electronic text editing was necessary. This was done to remove the non-prose content of the courses (notations, diagrams, graphics, tables, and complex typographic formats). Readability formulæ were devised to be used on continuous prose: and it is obvious that Open University course units contain varying amounts of non-prose content (in general, all forms of text contain more graphic and typographic elements than would have been the case half a century ago). Since the theoretical basis of the formulae was the linguistic structure of continuous prose, it is obvious that meaningless readings would be obtained from the non-prose sections. These were therefore edited out on our electronic files before sampling began.

In Flesch's instructions for sampling (Flesch 1951) he advocates taking 25-30 samples of 100 words each for a 'book'. This amounts to a sample of about 2% of the text. In our case we have taken about the same density of samples, but the sample length averaged over 1,000 words. The effect of this was to expand our samples to about 20% of the text (it was the availability of electronic files and the electronic application of the formula that made this possible). Therefore, we have every reason to believe that our samples were a fair and adequate representation of the prose element of the texts.

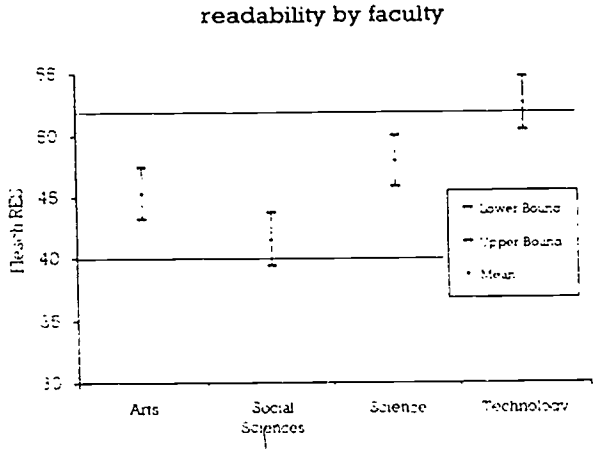


Figure 3: mean readability of the foundation course texts of four faculties.
The central points are mean of means. For A102 D103 and S102 this is the mean of means for units and for T102 it is the mean of means for topics. The upper and lower bounds mark the 95% confidence intervals.
The upper horizontal line in figure 3 corresponds to the readability of our test passage #2 (on which our sample did well) and the lower line corresponds to the readability of our passage 3 (on which 95% of our sample did not do well)

Comparisons between faculties

There are striking differences between the mean readability scores of the four faculties (Figure 2). If we ask whether faculties pitch their prose at the appropriate level for their students, then in the case of Technology the answer is broadly, yes. Also, the Faculty of Science could expect, with perhaps a more targeted editing effort, to adjust their text until it also matched the students' capabilities.

The Arts Faculty prose is overall significantly more difficult than the Technology faculty. Some allowance can be made for the fact that in our test sample the mean scores for Arts students were the best of the four faculties, and the Technology students the lowest. However these differences were not so large as the differences in mean readability of the course units. One can conclude that the Arts foundation course is not only somewhat overloaded, but is demanding in the level of reading skill required of students.

Lastly, we come to the Social Sciences foundation course, where the overall difficulty of the prose in the course units will without question cause difficulty for many of their students. However, as we saw above, the Social Sciences faculty have been more successful than Arts in controlling the overall text workload on the course. That is a counterbalancing factor.

Variation between course units

Of course, the faculty means mask great variety at the course unit level. Every faculty has some units (or topics) which are written in entirely suitable prose for their students, and every faculty has material which is written in prose which demands reading skills which we now know most of our students do not have.

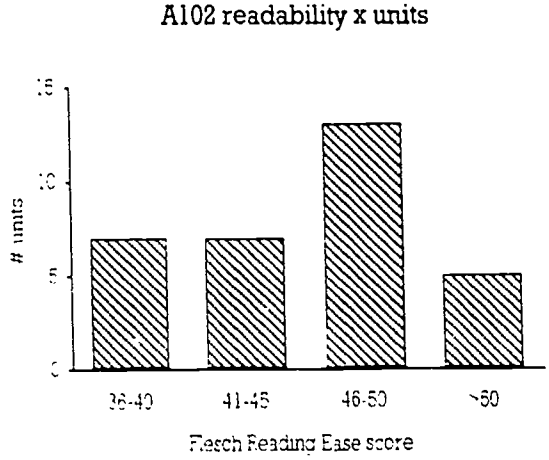
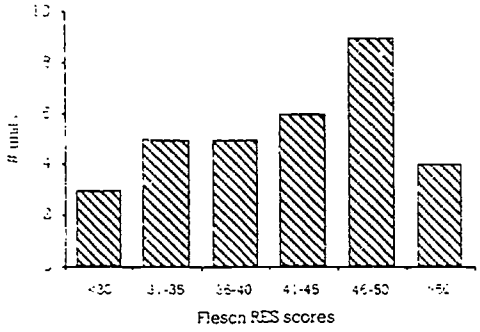
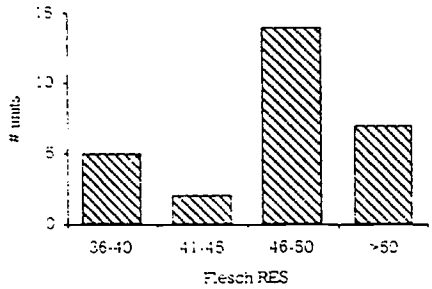


Figure 4: Variation within faculties at course unit or topic level (continued over).

D103 readability x units



S102 readability x units



T102 readability x topics

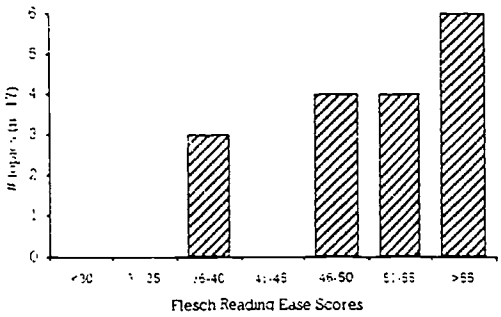


Figure 4 Variation within faculties at course unit or topic level.

Variation between blocks or disciplines

Different courses group their units by various principles; two of which stand out. One can group by discipline or quasi-discipline (S103, D103, A102) or one can group by physical binding ('the block').

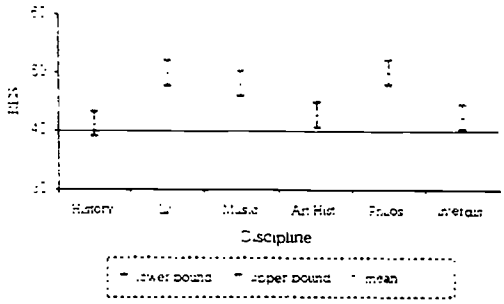
T102 is composed of blocks which have a highly complex series of threads or topics running through them

Variation between authors

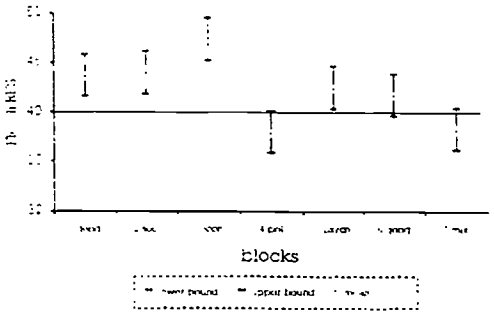
We have data on individual authors for those courses which identify the authors. A striking finding is the narrow range of scores for each author (as summarised by the standard deviations and 95% confidence intervals) compared to the range of mean scores for the whole group of authors.

This means that to a large extent readability scores are characteristic of the individual and remain fairly constant even when individuals write on different topics.

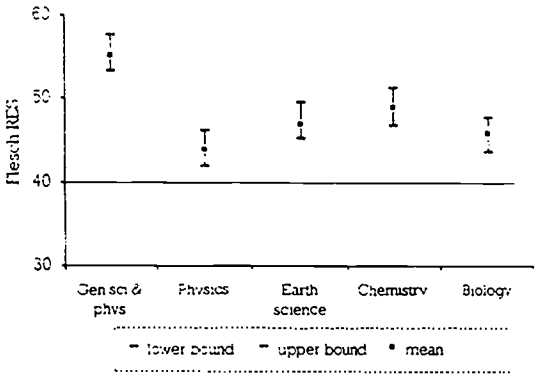
A102 readability x discipline



D103 readability x block



S102 readability x discipline



T102 readability x block

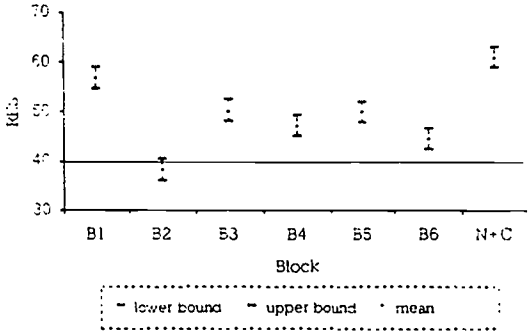


Figure 5 Variation between blocks or disciplines
The apparently greater variability within D10? blocks compared to the other faculties is an artefact of the graphical scaling: the summary data is in Appendix 2
S102 and D103 have blocks which are discipline-based where T102 has blocks which are theme-based
In T102 the numeracy and computing sections have been taken out of blocks and presented separately (N+C) The aberrant block 2 is a 1996 replacement

A102 readability by authors

Author	samples	mean RES	SD	95% ci
B Strang	1	57	–	–
G Martin	7	55	2 4	52 0–56.7
C.Cunningham	3	53	1 4	50 9–55 1
A.Calder		53	–	–
R Middleton	4	52	1 7	49 4–53 6
B.Fer	4	50	1.7	48 1–52.4
S Brown	12	47	3 3	45.3–49 6
D.Burrows	12	45	3 5	43.2–47.4
D.Walder	7	45	2.4	42.7–47.0
S.Meikle	2	44	1	41.9–46.1
E.Langmuir	12	43	3.3	40.6–44 9
J Golby	5	43	2	41.3–45 6
A.Marwick	14	42	3 6	39.7–43.9
G.Perry	1	40	–	–
G Roberts	2	38	1	36.4–40.6
M.Bartholomew	3	38	1 4	35.9–40 1
J.Bellamy	1	37	–	–
B.Purdue	8	36	2 6	33.9–38 1
G Parsons	5	29	2	27.1–31.3

Mean of means. 44.2
SD=4.2: 95% confidence interval = 42.4–46 7

In the A102 data the total range of author means is 28 points on the Flesch RES. yet the maximum standard deviation of any individual author is only 3.6. To put it another way. the best predictor of the readability of a section of course text is the identity of its author.

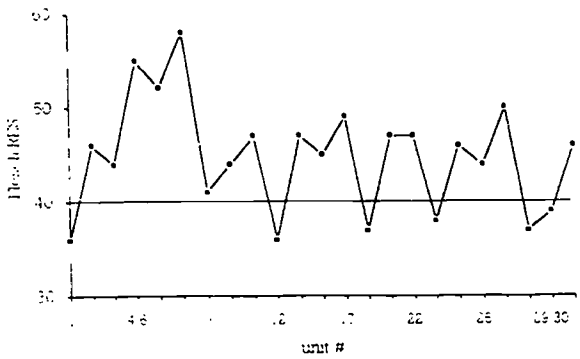
Many of the questions one would like to ask have not been investigated. For example, we do not know much from previous research about the extent to which individuals can adapt their writing style; there is likely to be much variation in this also. At any rate. over the production period of a foundation course our authors are remarkably consistent in the difficulty of the prose they turn out.

D103 readability by authors

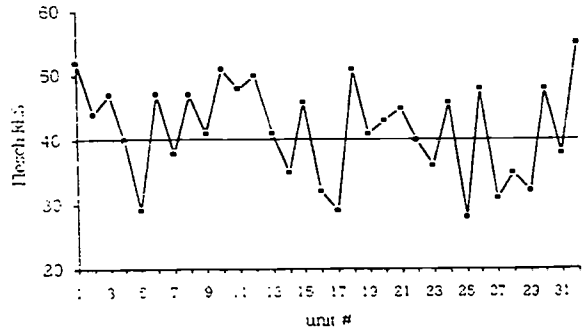
Author	samples	mean RES	SD	95% ci
D.Massey	8	48.5	2.6	46.4–50.6
E.Storkey	7	47 1	2 4	45 0–49.3
D.Coates	15	46 4	3 7	44 3–48.5
S Hall	5	46 4	2	44 3–48.5
J Allen	9	46 1	2 8	44.0–48 2
V Brown	7	45	2 4	42.9–47.1
M.Wetherell	10	44 2	3	42 1–46.3
Watson/Costello	7	43.6	2 4	41.4–45.7
B Bocock	3	40 7	1 4	38.5–42.8
R.Stevens	8	40.2	2 6	38.1–42.4
R.Maidment	4	39.2	1.7	37.1–41 4
K Thompson	3	38 3	1 4	36 2–40 5
D Deacon	2	36	1	33.9–38 1
O Hartley	2	34 5	1	32.4–36.6
D Denver	2	34 5	1	32.4–36 6
P Bradshaw	4	32 5	1 7	30.4–34 6
G McLennan	6	32	2 2	29 9–34 1
R Locker	2	27 5	1	25 4–29 6

Mean of means = 39.9 SD = 4 2 95% conf. int. = 37 7–42 0

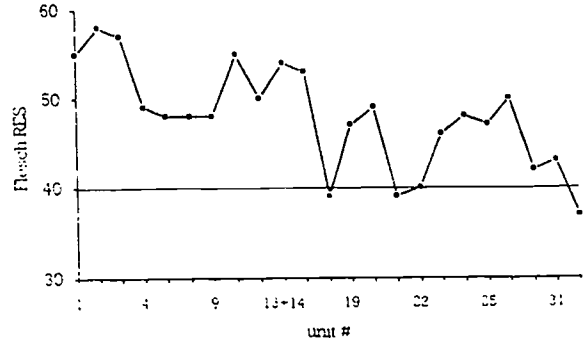
A102 readability x unit order



D103 readability x unit order



S102 readability x unit order



T102 topics in order

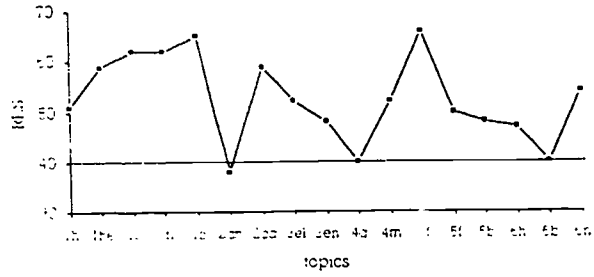


Figure 6 Variation in course unit readability in sequence from first to last unit in course

The D103 author data is similar: the range of mean RES between authors is 21 whereas the highest standard deviation of individual authors is 3.7.

Variation in units by sequence

In 1995 D103 tutor-counsellor Roger Fry carried out a survey of course readability using Flesch on small samples of 600 words per block (Fry, personal communication). He concluded that there was no progression in readability through the course, a finding which, with our much more extensive sampling, we can confirm (unit order charts, figure 5, previous page).

By way of contrast, there is clearly progression in S102, no doubt as a result of deliberate course team policy. Of course, it is an open question as to whether there ought to be such progression in reading difficulty. Conceptual difficulty will increase as the course proceeds, especially in subjects that have an hierarchical structure, but ought not the prose maintain a level which allows most of the students to engage in the concepts rather than struggle with the English?

Mathematics foundation courses

We did not include M101 in our survey because in 1996 its life comes to an end. The Mathematics and Computing Faculty are creating a suite of three half-credit courses to take its place, and we will be estimating the readability of prose on the level 1 half-credit MU120 (Open Mathematics) during 1996.

In general, it is valid to use readability measures on elementary maths courses since they do contain significant amounts of prose.

Higher level courses

We may reasonably expect higher level courses to exhibit more difficult prose if for no other reason than the general lack of support for the beleaguered author. These course teams are smaller, so each academic does more writing, and the constellation of colleagues, editors and educational technologists which surround foundation courses are absent or scarce in a higher level course. We might also expect more variation between courses, remembering our key finding that the best predictor of the readability of prose is the author's identity.

At present we have data on one higher level course, A317 Themes in British and American History: a comparative approach, c.1760–1870 [the title itself is quite stretching!]. The horizontal rule at 40 RES marks the difficulty of passage 3 in our reading test. Most of the prose is quite difficult, and some of it is extremely difficult and bound to cause problems for most students.

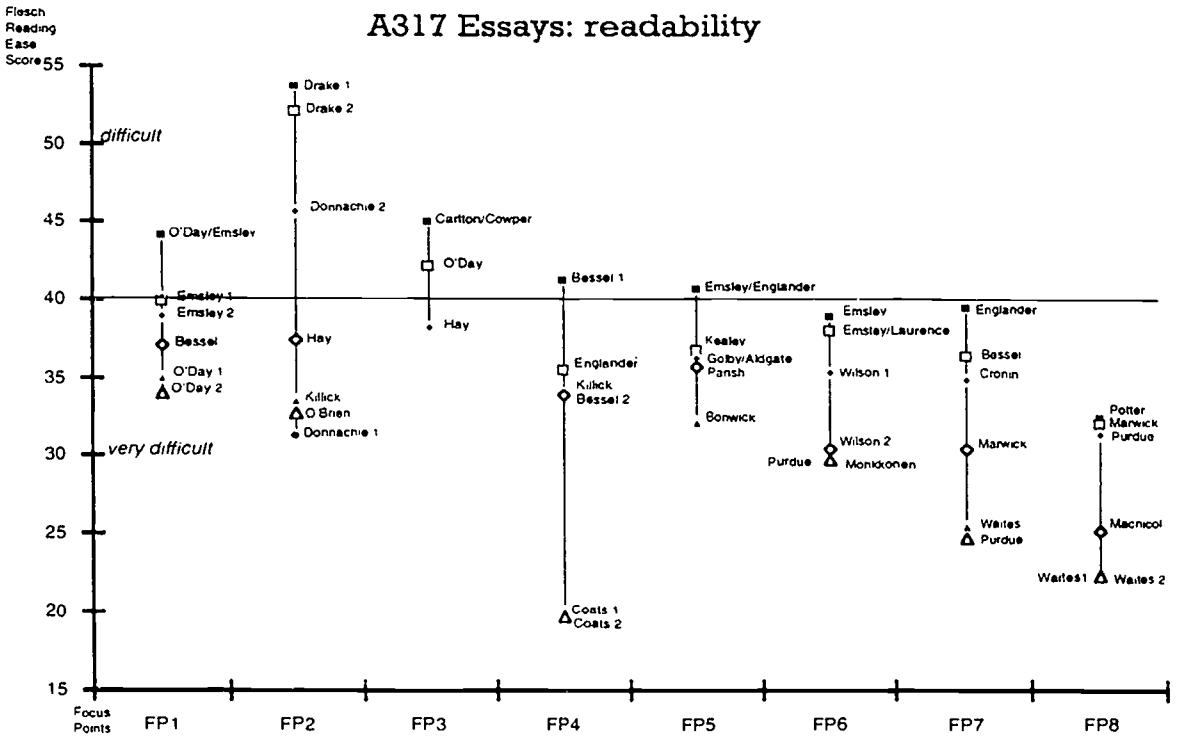


Figure 7. The readability of A317 in topic sequence together with data points for individual authors. Data and figure courtesy of Clive Lawless.

Conclusion

Not only do our students have widely varying reading skills, but our foundation courses contain text which varies greatly in its readability. Much of the prose is certain to cause difficulty with students who score in the bottom half of our reading tests. We already know that such students are more likely than average to drop out before final registration, and our hypothesis is that the difficulty and sheer quantity of prose in the courses is one main reason for this.

Data on readability has not been available previously to course teams, so perhaps it is no surprise that the means of course units may show a spread of up to thirty points on the Flesch RES scale. This speaks of an absence of monitoring which might be corrected in future. Perhaps more surprising is that individual authors are remarkably consistent in the prose they turn out, even when they write on different topics and units. It will be interesting to see how flexible they are when presented with this kind of data: one expects the better writers to be more flexible than the less expert ones. At a course teams's disposal is a variety of editing skills which need targeting to be effective.

For those who like factual nuggets to take away and treasure, it is clear that the single best predictor of the readability of course prose is the identity of its author.

Although we have collective responsibility for our courses, writing is a personal activity. If the prose is difficult then there is at present no formal evidence that later processes improve it. Yet a great deal of work goes into the discussion, revision and editing of the course texts. It can be argued that the prose would be in much worse shape if not for the efforts of editors and others, and no doubt there is some truth in that.

The story of this project is not yet complete. We now have the data on assignment scores, and the final examination results for the 1995 foundation courses. The next report will analyse this data, and discuss its relation to our model.

Postscript

The readability of the prose in this report is Flesch RES 46. To calculate this value text other than continuous prose (such as graphics, tables, lists, headings, captions and references) was removed. This left about 3,250 words of continuous prose. The Flesch Reading Ease Score was got by using the Grammar tool in Microsoft Word after the grammar routines were disabled in Preferences.

References

Carver RP Reading rate: a review of research and theory. New York: Academic Press 1990

Flesch R How to test readability. NY: Harper 1951.

Hirsch ED Jr The Philosophy of Composition. Chicago IL: The University of Chicago Press 1977.

Johnson KH, Relova RP Jr, and Stafford JP An analysis of the relationship between readability of Air Force procedural manuals and discrepancies involving non-compliance with the procedures. (Document AD 750 917. National Technical Information Service) Air Force Institute of Technology, Air University 1972.

Kern RP, Sticht TG and Fox LC Readability, reading ability and readership (Professional Paper 17-20) Alexandria VA: Human Resources Research Organisation 1970.

Klare GR The measurement of readability. Ames: Iowa State University Press 1963.

Klare GR Assessing readability. Reading Research Quarterly 10, 62-102, 1974-75.

Klare GR Readability in Pearson PD (ed) Handbook of Reading Research. New York: Longman 1984.

Klare GR and Smart K Analysis of the readability level of selected United States Armed Forces Institute printed instructional materials. Journal of Educational Research 67, 176, 1973.

Macdonald-Ross M Language in texts in Lee S Shulman (ed) Review of Research in Education 6. 229-275 Itasca IL: Peacock 1979.

Macdonald-Ross M and Scott B Results of the survey of OU students' reading skills. Technical Report #3. Text and Readers Programme, IET, The Open University 1995

Macdonald-Ross M and Scott B OU students reading skills and final registration. Technical Report #4. Text and Readers Programme, IET, The Open University 1995.

Murphy DR Test proves short words and sentences get best readership. Printer's Ink 218, 61-64, 1947.

Murphy DR How plain talk increases readership 45 per cent to 66 per cent. Printer's Ink 220, 35-37, 1947.

Scott B and Macdonald-Ross M A postal survey of OU students' reading skills. Technical Report #3. Text and Readers Programme IET The Open University 1995

Swanson CE Readability and readership: a controlled experiment. Journalism Quarterly 25, 339-343, 1948

Appendix 1: A102 data

A 102 Textload

There are 34 weeks: 1 week is set aside for Easter. 1 week is set aside for Summer School. The OU assumes a typical week requires 12 hours of study

A102 Preparatory Pack (c45,100w)

A102 has 32 units, each of which is part of one week's study. They are multiply authored, there are a total of 19 authors.

Course Guide (c19 pages, c500wpp)

Glossary (c20p, c500wpp)

Chronology (c1000 w)

Illustration booklet (119 plates index c3000w)

Broadcast Notes (c120 p, c400wpp) and Cassette Notes (c30p, c400wpp).

Set books: Best G Mid-Victorian Britain (c320p, c350wpp)

Dickens C Hard Times with intro and notes (c320p, c380wpp).

Gombrich EH Art and Illusion (c300p, c500 wpp) 50% or more of this text needs to be read

Course Reader: Golby J (ed) Culture and Society in Britain 1850-90: a source book of contemporary writings. (c330p, c450wpp)

Assignments Booklet (c12p, c400wpp)

Supplementary Material Booklet (c48p, c400wpp).

Hard Times Study Guide (c10p, c400wpp)

A102 Literature Supplement (c12p, c400wpp)

Library Guide (c14p, c400wpp).

A102 Summer School booklet: students' guide (8p, c400wpp).

Summer School Texts for project work (varous).

A102 Summer School Computing Experience: Students' Guide (12p, c400wpp).

A102 After the Summer School (5p, c400wpp).

A102 Rationale for the Examination (3p, c300wpp) and A102 Preparing for the Examination (6p, c500wpp): specimen examination paper (6p, c500wpp)

32 tv programmes of 50 minutes each. With preparation and follow-up work study time is expected to be about 2 hours per programme.

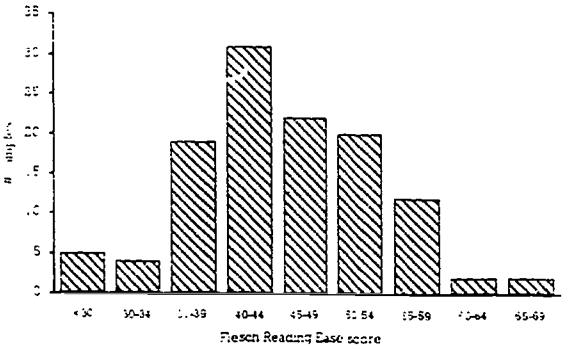
There are 16 radio programmes of 20 minutes each study time for each expected to be 30 mins.

5 audio-cassettes (4 at 90 mins, one at 30 mins)

An estimation of total reading required by number of words (Preparatory Pack and Summer School Excusal Pack not included).

Units	512,200
Course Guide	9,800
Chronology	1,000
Glossary	10,000
Broadcast Notes	50,600
Cassette Notes	16,800
Set books:	
Best	112,800
Dickens	112,000
Gombrich (50%)	75,000
Reader Golby (90%)	137,000
Assignments booklet	4,800
Supplementary material	33,000
Hard Times study guide	4,000
Library Guide	5,900
SS texts (80%)	81,900
SS guide	3,200
After SS	2,000
Stop Press	3,500
Total	1,175,300 words

A102 readability



For each foundation course the readability of over 100 samples, each of one to two thousand words, is the core data.

This core data is then analysed into unit, block, discipline, author and unit sequence in the main body of this report.

A102: Readability of Course Units

Unit	samples	mean RES	SD	95% ci
1	4	36	1.7	34.9–39.1
2	3	46	1.4	43.9–48.1
3	4	44	1.7	41.9–46.1
4-6	7	55	2.5	52.0–56.3
7	4	52	1.7	49.4–53.6
8	3	58	1.4	56.2–60.5
9	6	41	2.2	40.0–44.3
10	5	44	2.0	42.5–46.7
11	4	47	1.7	43.4–47.6
12	3	36	1.4	33.9–38.1
13-15	12	50	3.3	48.0–52.2
16	3	45	1.4	41.5–45.8
17	6	49	2.2	46.9–51.1
18-19	8	37	2.6	34.5–38.8
20-21	7	47	2.4	44.4–48.7
22	4	47	1.7	45.9–50.1
23-24	5	38	2.0	35.7–39.9
25	4	46	1.7	43.1–47.4
26	6	44	2.2	42.7–46.9
27	3	50	1.4	47.5–51.8
28	4	37	1.7	34.6–38.9
29-30	7	39	2.4	36.6–40.8
31-32	5	46	2.0	43.5–47.7

Mean of means = 45.3 SD = 5.6

95% confidence interval = 43.2–47.4

A102: Readability by discipline

Discipline	samples	mean RES	SD	95%ci
History	18	41.1	4.1	39.0–43.2
Literature	13	49.8	3.5	47.7–52.0
Music	17	47.9	4.0	45.8–50.1
Art History	12	42.7	3.3	40.6–44.9
Philosophy	15	50	3.7	47.9–52.1
Interdisc.	41	42.4	6.3	40.2–44.5

Mean of means 45.6 SD = 3.2

95% confidence interval = 43.5–47.8.

Key to Units

History	Units 1-3, 16, 28
Literature	Units 4-6, 26
Music	Units 7-9, 25 Art History Units 10-12
Philosophy	Units 13-15, 27
Interdisc.	Units 17-24, 29-32

Appendix 2: D103 data

D103 Textload

There are 34 weeks: 1 week is set aside for Easter. 1 week is set aside for Summer School. The OU assumes a typical week requires 12 hours of study

D103 has 32 Units. Each is part (Course Guide says roughly half) of one week's study. They are multiply authored: 23 authors in toto. Units have sections on study skills. There are 8 of these (3 in Block 1, 1 in Block 2, 1 in Block 3, 1 in Block 5, 1 in Block 6 and 1 in Block 7).

Course Guide (c17 pages, c500wpp)

Glossary Index (c31 pages, c500 wpp)

Media Booklet (c76 pages, c350 wpp)

Set book, *The Good Study Guide* (not all has to be read: c206 pages, c350 wpp).

Reader with 22 chapters to be read, *Society and Social Science: A Reader* (c300 pages, c500 wpp)

Assignments booklets (c17 pages, c600 wpp)

Booklet, *After D103: OU Courses in Social Science* (8 pages, c750 wpp)

Booklet *Preparing for D103 Summer School* (8 pages, c600 wpp)

Summer School set of workbooks (c65 pages of text, c450 wpp)

For students not attending Summer School there is a *Summer School Excusal Pack* (c35 pages, c250 wpp)

Stop Press messages with news, notes for radio programmes and errata (14 pages, c250 wpp)

D103 Preparatory Pack for use before starting the cours proper.

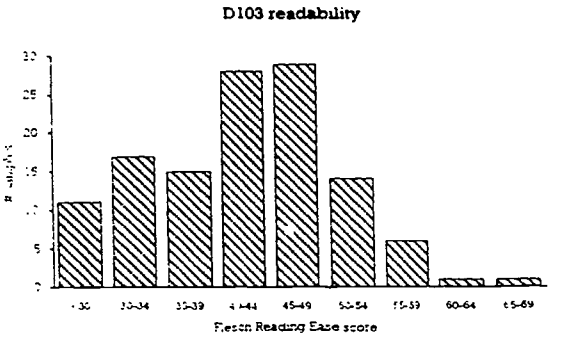
16 tv programmes; each is 50 minutes. With preparation and follow-up work, study time is expected to be about 2 hours per programme.

16 radio programmes; each is 20 minutes, total study time is expected to be 30 minutes.

7 audio-cassettes, one for each Block. Each holds a maximum of one hour's listening; extra study time is allocated. Audio-cassette 7 is accompanied by a booklet, *D103 Endnotes for Block 7* (c15 pages, c300 wpp).

An estimation of total reading required by number of words (Preparatory Pack and Summer School Excusal Pack not included).

Units:	574,000
Course Guide	8,500
Glossary	14,500
Media Booklet	26,000
Set Book	72,100
Reader	150,000
TMA Booklets	10,000
After D103	6,000
SS Workbooks	29,250
Stop Press	3,500
Endnotes	4,500
Total	897,550 words



Distribution of D103 samples

D103: readability of course units

Unit	# samples	mean RES	SD	95% ci
1	4	52	1.7	49.6–53.9
2	4	44	1.7	42.4–46.6
3	4	47	1.7	43.9–48.1
4	3	40	1.4	37.9–42.1
5	2	29	1	25.9–30.1
6	4	47	1.7	44.6–48.9
7	3	38	1.4	36.2–40.5
8	4	47	1.7	44.4–48.6
9	2	41	1	38.9–43.1
10	4	51	1.7	49.1–53.4
11	4	48	1.7	45.1–49.4
12	3	47	1.4	45.5–49.8
13	2	41	1	38.9–43.1
14	3	35	1.4	32.9–37.1
15	4	46	1.7	44.1–48.4
16	4	32	1.7	30.4–34.6
17	4	28	1.7	25.4–29.6
18	2	51	1	48.4–52.6
19	5	41	2	39.1–43.3
20	4	43	1.7	40.1–45.1
21	5	45	2	42.8–47.1
22	3	40	1.4	36.5–40.8
23	4	45	1.7	43.1–47.4
24	4	46	1.7	43.1–47.4
25	4	28	1.7	26.1–30.4
26	4	48	1.7	45.4–49.6
27	2	26	1	24.4–28.6
28	4	35	1.7	33.1–37.4
29	5	32	2	29.3–33.5
30	3	48	1.4	45.9–50.1
31	4	38	1.7	36.4–40.1
32	3	51	1.4	48.2–52.5

Mean of means = 41.6 SD=5.6

95%confidence interval = 39.4–43.7

D103: readability of course blocks

Block	samples	mean	SD	95% ci
1	18	43.6	4.1	41.6–45.8
2	14	43.9	3.6	41.8–46.1
3	14	47.4	3.6	45.2–49.5
4	18	37.9	4.1	35.8–40.1
5	18	42.4	4.1	40.3–44.6
6	17	41.7	4.0	39.6–43.8
7	22	38.2	4.6	36.1–40.4

Mean of means = 42.2 SD = 2.4

95% confidence interval = 40.0 –44.3

Appendix 3: S102 data

S102 Textload

34 weeks: 1 week is set aside for Easter, 1 week is set aside for Summer School. The OU assumes a typical week requires 12 hours of study.

S102 has 32 Units. Each is part of one week's study (c. 1500 pages, c. 450 words per page).

Introduction and Study Guide (c. 30 pages, c 350 words per page).

Study Comments (28 pages, c. 250 words per page).

Course Index (c. 35 pages, c. 600 words per page). This is provided for occasional use, for when students wish to locate a particular topic within the course texts.

Glossary (c. 55 pages, c. 600 words per page).

Assignments Booklets (c. 65 pages, c. 450 words per page).

Introduction to Summer School (c.4 pages, c. 300 words per page).

Summer School Laboratory Notebooks (c 90 pages, c. 300 words per page).

Revision Guide (c. 8 pages, c. 450 words per page).

Specimen Examination Paper (c. 20 pages, c. 600 words per page).

Stop Press messages with news, notes and errata (c. 10 pages, c.250 words per page).

35 tv programmes; each is 25 minutes. Notes on the content of the programmes are included in the relevant Course Unit.

5 audio-cassettes (90 minutes each); total study time is expected to be longer than this. The tapes are used with other material in the Course Units to form audio-visual sequences, with activities for the student to carry out.

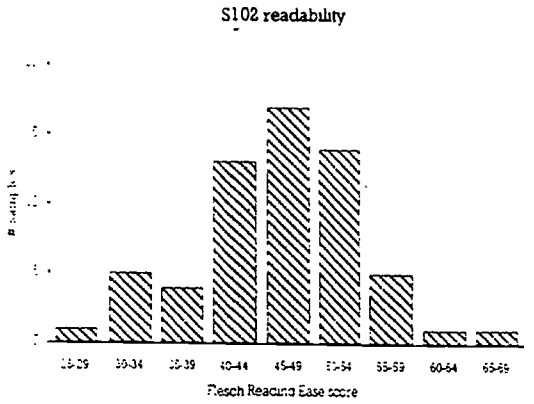
An *Experiment Kit* is provided for carrying out experiments. Short booklets with the kits give instructions for setting up apparatus and provide advice on safety.

Three Computer-Assisted Learning (CAL) programmes.

Preparatory materials, *Into Science*, for use before starting the course proper (c.220 pages, c. 400 words per page).

Estimation of total reading required by number of words (preparatory material, *Experiment Kit* booklets and *Course Index* not included):

Units	675000
Introduction and Study Guide	10500
Study Comments	7000
Glossary	33000
Assignments Booklets	29200
Introduction to Summer School	1200
Summer School Lab. Notebooks	27000
Revision Guide	3600
Specimen Examination Paper	12000
Stop Press	2500
Total	801000



Distribution of S102 samples

S102: readability of course units

Unit	# samples	mean RES	SD	95% ci
1	4	55	1.7	52.9–57.1
2	3	58	1.4	56.2–60.5
3	3	57	1.4	55.2–59.5
4	2	49	1.0	46.9–51.1
5-6	8	48	2.6	44.7–49.0
7-8	6	48	2.2	46.0–50.3
9	3	48	1.4	46.5–50.8
10	5	55	2.0	52.5–56.7
11-12	5	50	2.0	48.5–52.7
13-14	5	[illegible]	2.0	52.3–56.5
15	3	53	1.4	50.5–54.8
17-18	5	39	2.0	37.9–42.1
19	2	47	1.0	44.4–48.6
20	2	49	1.0	46.4–50.6
21	2	39	1.0	37.4–41.6
22	1	40	—	—
23	2	46	1.0	44.4–48.6
24	3	48	1.4	46.2–50.5
25	2	47	1.0	44.4–48.6
26	1	50	—	—
30	3	42	1.4	40.5–44.8
31	3	43	1.4	40.5–44.8
32	4	37	1.7	34.4–38.6

Mean of means = 47.9. SD = 4.7

95% confidence interval = 45.8–50.0

S102: readability by discipline

	samples	m. RES	SD	95% ci
Gen. sci	12	55	3.3	53.3–57.5
Physics	18	44	4.1	42.0–46.2
Earth Sci	14	47	3.6	45.3–49.6
Chem	18	49	4.1	46.9–51.2
Biology	14	46	3.6	43.6–47.8

Mean of means = 48.2. SD = 2

95% confidence interval = 46.1–50.3.

General Science and Phys:	Units 1-4
Physics:	Units 9-10, 30-32
Earth Sciences:	Units 5-8, 27-29
Chemistry:	Units 11-18
Biology:	Units 19-26

Appendix 4: T102 data

T102 Textload

There are 34 weeks: 1 week is set aside for Easter. 1 week is set aside for Summer School. The OU assumes a typical week requires 12 hours of study.

T102 has 6 *Blocks*. Each *Block* has one mainstream topic and two, three or four tributary topics (c998 pages, c500wpp).

Introduction (c36 pages, c500wpp)

Introductory Supplement & Reference Book (c6 pages, c500wpp).

Supplementary Material in each block includes study guides, information about assignments, visual materials that accompany audiotapes, selected readings and miscellaneous items of information (c125 pages, c500wpp).

Software Guide (c10 pages, c300wpp). This is a reference document, to be consulted as required.

Communications Guide (c30 pages, c300wpp)

Summer School Handbook (c90 pages, c500wpp).

Specimen Examination Paper (c15 pages, c500wpp)

Preparing for Living with Technology and associated *Supplementary Material* for use before starting the course proper (c70 pages, c500wpp).

Working with Windows, 1, 2 and 3 to be studied before the course begins and for later reference (c160 pages, c300wpp). The guides are accompanied by 3 audio-cassettes (90 minutes each) and an applications disc. The software supports word-processing, spreadsheets and communications. There are also Computer Aided Learning (CAL) materials that support *Numeracy* (Blocks 1, 2, 3 & 6) and *Working with Chemical Symbols* (Block 4).

8 tv programmes, each of 50 minutes. They are designed to be studied in conjunction with study guides and Block texts. It is recommended that they are recorded for repeated viewing and note-taking.

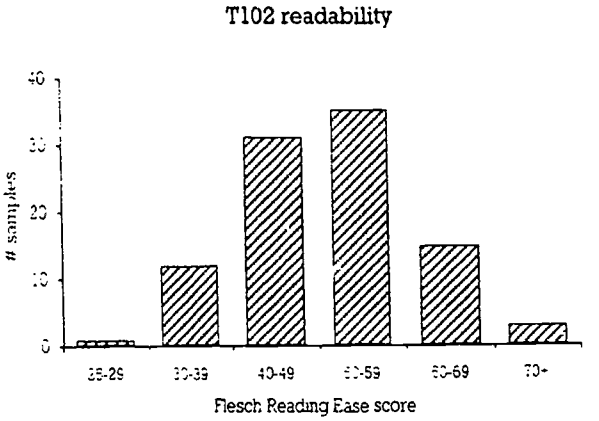
5 audio-cassettes (90 minutes each); total study time is expected to be longer than this. The tapes are used with other material in the Blocks to form audio-visual sequences, with activities for the student to carry out.

An estimation of total reading required by number of words (preparatory material and *Software Guide* not included):

Introduction	18,000
Blocks:	497,500
Introductory Supplement	3,000
Supplementary Material	62,500
Communications Guide	9,000
Summer School Handbook	45,000
Specimen Examination Paper	7,500
Total	642,500

The most obvious difference from A102 and D103 is the absence of Set Books or Readers: this greatly reduces the reading load.

T102: readability of some course topics



Distribution of T102 samples

BLOCK 1

Topic	samples	mean RES	SD	95% ci
Home	10	51	3.0	49.6–53.8
Heat	5	59	2.0	56.5–60.7
Str.& Mat.	5	62	2.0	59.7–63.9
Numeracy	6	62	2.2	59.2–63.5
Computing	5	65	2.0	63.3–67.5

BLOCK 2

Topic	samples	mean RES	SD	95% ci
Comm.	9	38	2.8	36.2–40.5
Computing	2	59	1.0	56.9–61.1

BLOCK 3

Topic	samples	mean RES	SD	95% ci
Electricity	5	52	2.0	50.3–54.5
Energy	6	48	2.2	46.4–50.6

BLOCK 4

Topic	samples	mean RES	SD	95% ci
Growth	5	40	2.0	38.1–42.3
Materials	8	52	2.6	49.5–53.7
Numeracy	4	66	1.7	63.4–67.6

BLOCK 5

Topic	samples	mean RES	SD	95% ci
Food	12	50	3.3	48.5–52.8
Biology	4	48	1.7	45.9–50.1

BLOCK 6

Topic	samples	mean RES	SD	95% ci
Health	9	47	2.8	44.3–48.6
Biology	4	40	1.7	38.4–42.6
Numeracy	5	54	2.0	51.9–56.1

Mean of means of blocks = 52.5. SD = 4.0
95% confidence interval = 50.4–54.7

Text & Readers Programme Publications

Technical Reports

1: The revolution in print technology
2: A postal survey of OU students' reading skills
3: Results of the survey of OU students' reading skills
4: OU students' reading skills and final registration
5: The readability of OU foundation courses

Newsflash

1: OU students' reading skills
2: Readability of OU courses

Michael Macdonald-Ross and Bernard Scott

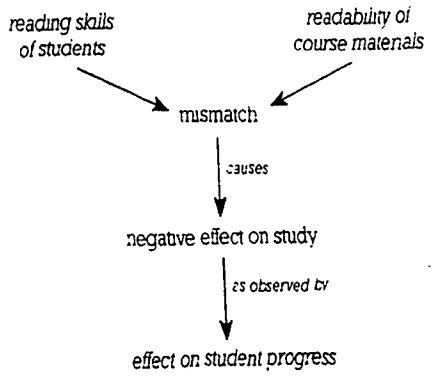
The readability of OU foundation courses

From our previous work on reading skills we know that many students find academic prose difficult. We now have data on the readability of foundation courses: such data could help courses better match the students' reading ability.

We think that mismatch between the reading skills of the students and the readability of course materials is an important cause of student drop-out and failure.

To test this idea we compare our data on student reading skills and text readability levels to data on the progress of our cohort throughout the year. So far the model has worked well.

Of course other factors do affect student progress. How reading mismatch stands relative to other factors in student progress is being investigated.



Readability measures

Readability measures predict the difficulty of prose: usually they combine an index of syntactic difficulty (eg sentence length) with an index of vocabulary (eg word length). The most frequent words in a language are the best known, and from the work of George Zipf in the 1930s we know that there is a relationship between the length of words and their frequency in use.

For our work we used an electronic version of the Flesch Reading Ease Score which is offered as a tool in Microsoft Word under the Grammar facility.

Results

The chart 'readability by faculty' (above right) shows the mean readability of four foundation courses, and illustrates the variation in their mean prose difficulty.

The upper horizontal line corresponds to the readability of our reading test passage 2 (on which most students did well), the lower line corresponds to the readability of our test passage 3 (on which 95% of our sample did not do well).

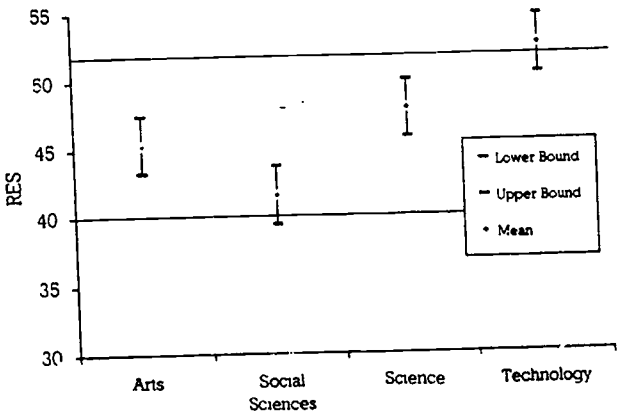
Faculty of Technology has put in a great effort to make T102 accessible, and deserves congratulation on its readability. The effort was needed because, as it happens, technology entrants scored lowest in our reading skill tests.

Each course can be analysed into blocks, units, disciplines, authors &c. The samples for each course number over a hundred (eg 122 samples for D103); the length of samples averages well over a thousand words each. This probably makes it the most substantial readability sample of university-level teaching material ever conducted.

Above right, S102 is analysed into disciplines (apart from the first data set, the sequence is not in course unit order). From this kind of data course team management get hints on which parts of a course might need priority attention during revision and editing.

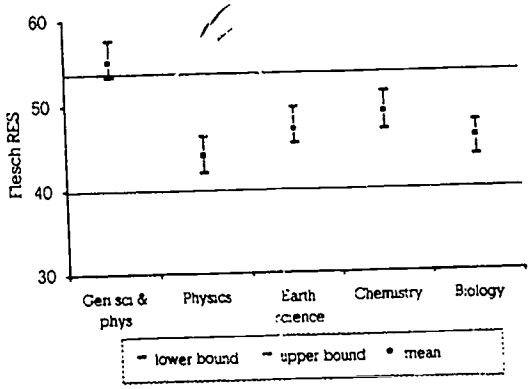
In D103 the blocks are mostly built around disciplines. A picture emerges that blocks four and seven have significantly more difficult prose than the rest: our idea is that in future this kind of data should be picked up during course development.

readability by faculty

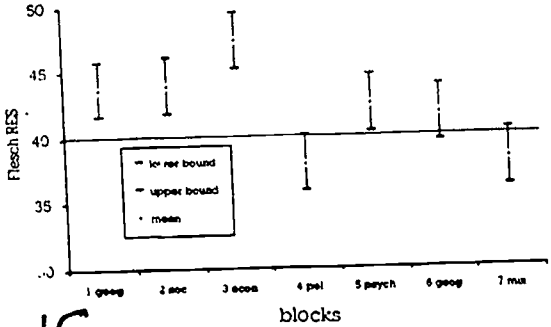


M101 is in its last year, and was not included in this study
The readability of MU120 will be examined later this year.

S102 readability x discipline



D103 readability x block



A point made by a D103 tutor-counsellor was that the course showed no progression in difficulty of prose when analysed by unit order. By contrast S102 seems to have achieved this quite well.

S102 readability x unit order

Opposite is a table which shows that ten authors in a course produce prose at least as difficult as passage 3 in our test; of those ten, six write prose so difficult that it is bound to be an obstacle to comprehension and student progress.

The range of means for authors in a course is wide (20 to 30 points on the Flesch RES scale) whereas the range of standard deviations for individual authors is quite narrow (about 3 points on the scale). Thus the best predictor of the readability of a course unit is the identity of the author.

A list of authors in order of readability is no comment on their quality as expert academics. Of course, readability is not a measure of the academic worth of the course material: but on the other hand students cannot benefit from erudition unless they can read and understand it!

We developed the term 'textload' for the minimum amount of text in words which a student would need to read to comply with study guides in a course. The variation between faculties is considerable. If D103 is about the right textload (we think it is) then A102 is overloaded. This will put extra pressure on students with sub-standard reading skills.

S102 and T102 have less text because they do not use set books and readers at foundation course level. Instead, they use home experimental kits and computers to allow for practical activities associated with the course. Therefore, in S102 and T102 the textload is only a partial guide to the overall workload. Of course, all courses have assignments and other activities which compete for scarce time. Assignments can be very time-consuming for some students.

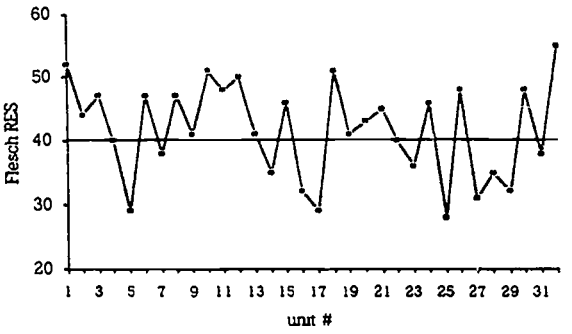
Higher-level courses are likely to have problems, since they do not get the attention lavished on foundation courses.

A317 is a case in point; notice that much more than half the material is more difficult than our passage 3 test. Of course, many of the weaker readers will have dropped out before reaching third level. This illustrates the tension between the University's wish to be open, and the consequent fact that many students' reading standard is closer to the Daily Mail than to our academic prose.

A full account of this research is available: ***The readability of OU foundation courses****, IET's Text & Readers Programme Technical Report #5.*

Data+chart courtesy of Clive Lawless

D103 unit readability in order

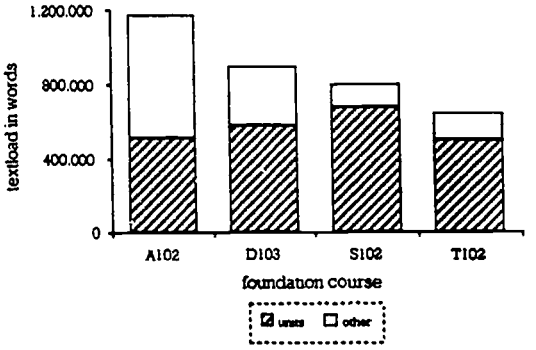


D103: Course unit authors in order of readability

	# samples	mean RES	SD	95% confidence int.
Author #1	8	48.5	2.6	46.4-50.6
Author #2	7	47.1	2.4	45.0-49.3
Author #3	15	46.4	3.7	44.3-48.5
Author #4	5	46.4	2	44.3-48.5
Author #5	9	46.1	2.8	44.0-48.2
Author #6	7	45	2.4	42.9-47.1
Author #7	10	44.2	3	42.1-46.3
Authors #8+9	7	43.6	2.4	41.4-45.7
Author #10	3	40.7	1.4	38.5-42.8
Author #11	8	40.2	2.6	38.1-42.4
Author #12	4	39.2	1.7	37.1-41.4
Author #13	3	38.3	1.4	36.2-40.5
Author #14	2	36	1	33.9-38.1
Author #15	2	34.5	1	32.4-36.6
Author #16	2	34.5	1	32.4-36.6
Author #17	4	32.5	1.7	30.4-34.6
Author #18	6	32	2.2	29.9-34.1
Author #19	2	27.5	1	25.4-29.6

Mean of means = 39.9. SD = 4.2: 95% confidence interval = 37.7–42.0

textload x foundation course



A317 Essays: readability